SEQUENCE LISTING

<110> Hoechst Marion Roussel
Bordon-Pallier, F.
Rocher, C.

<120> Human htFIIIA gene and coded htFIIIA protein

<130> 146.1364

<140> US 09/831,426
<141> 2001-05-08

<160> 10

<170> PatentIn Vers. 2.0

<210> 1
<211> 1273
<212> DNA
<213> Human

<220>
<221> CDS
<222> (176)..(1270)

<400> 1
atgcgcagca gcggcgccga cgcggggcgg tgcctggtga ccgcgcgcgc tcccggaagt 60

gtgccggcgt cgcgcgaagg ttcagcaggg agccgtgggc cgggcgcgcc ggttcccggc 120

acgtgtctcg gcacgtggca gcgcgcctgg ccctgggctt ggaggcgccg gcgcc ctg 178
Met
1

gat ccg ccg gcc gtg gtc gcc gag tcg gtg tcg tcc ttg acc atc gcc 226
Asp Pro Pro Ala Val Val Ala Glu Ser Val Ser Ser Leu Thr Ile Ala
5 10 15

gac gcg ttc att gca gcc ggc gag agc tca gct ccg acc ccg ccg cgc 274
Asp Ala Phe Ile Ala Ala Gly Glu Ser Ser Ala Pro Thr Pro Pro Arg
20 25 30

ccc gcg ctt ccc agg agg ttc atc tgc tcc ttc cct gac tgc agc gcc 322
Pro Ala Leu Pro Arg Arg Phe Ile Cys Ser Phe Pro Asp Cys Ser Ala
35 40 45

aat tac agc aaa gcc tgg aag ctt gac gcg cac ctg tgc aag cac acg 370
Asn Tyr Ser Lys Ala Trp Lys Leu Asp Ala His Leu Cys Lys His Thr
50 55 60 65

ggg gag aga cca ttt gtt tgt gac tat gaa ggg tgt ggc aag gcc ttc 418
Gly Glu Arg Pro Phe Val Cys Asp Tyr Glu Gly Cys Gly Lys Ala Phe
70 75 80

atc agg gac tac cat ctg agc cgc cac att ctg act cac aca gga gaa 466

```
Ile Arg Asp Tyr His Leu Ser Arg His Ile Leu Thr His Thr Gly Glu
            85                  90                  95

aag ccg ttt gtt tgt gca gcc act ggc tgt gat caa aaa ttc aac aca   514
Lys Pro Phe Val Cys Ala Ala Thr Gly Cys Asp Gln Lys Phe Asn Thr
        100                 105                 110

aaa tca aac ttg aag aaa cat ttt gaa cgc aaa cat gaa aat caa caa   562
Lys Ser Asn Leu Lys Lys His Phe Glu Arg Lys His Glu Asn Gln Gln
    115                 120                 125

aaa caa tat ata tgc agt ttt gaa gac tgt aag aag acc ttt aag aaa   610
Lys Gln Tyr Ile Cys Ser Phe Glu Asp Cys Lys Lys Thr Phe Lys Lys
130                 135                 140                 145

cat cag cag ctg aaa atc cat cag tgc cag cat acc aat gaa cct cta   658
His Gln Gln Leu Lys Ile His Gln Cys Gln His Thr Asn Glu Pro Leu
                150                 155                 160

ttc aag tgt acc cag gaa gga tgt ggg aaa cac ttt gca tca ccc agc   706
Phe Lys Cys Thr Gln Glu Gly Cys Gly Lys His Phe Ala Ser Pro Ser
            165                 170                 175

aag ctg aaa cga cat gcc aag gcc cac gag ggc tat gta tgt caa aaa   754
Lys Leu Lys Arg His Ala Lys Ala His Glu Gly Tyr Val Cys Gln Lys
        180                 185                 190

gga tgt tcc ttt gtg gca aaa aca tgg acg gaa ctt ctg aaa cat gtg   802
Gly Cys Ser Phe Val Ala Lys Thr Trp Thr Glu Leu Leu Lys His Val
    195                 200                 205

aga gaa acc cat aaa gag gaa ata cta tgt gaa gta tgc cgg aaa aca   850
Arg Glu Thr His Lys Glu Glu Ile Leu Cys Glu Val Cys Arg Lys Thr
210                 215                 220                 225

ttt aaa cgc aaa gat tac ctt aag caa cac atg aaa act cat gcc cca   898
Phe Lys Arg Lys Asp Tyr Leu Lys Gln His Met Lys Thr His Ala Pro
                230                 235                 240

gaa agg gat gta tgt cgc tgt cca aga gaa ggc tgt gga aga acc tat   946
Glu Arg Asp Val Cys Arg Cys Pro Arg Glu Gly Cys Gly Arg Thr Tyr
            245                 250                 255

act act gtg ttt aat ctc caa agc cat atc ctc tcc ttc cat gag gaa   994
Thr Thr Val Phe Asn Leu Gln Ser His Ile Leu Ser Phe His Glu Glu
        260                 265                 270

agc cgc cct ttt gtg tgt gaa cat gct ggc tgt ggc aaa aca ttt gca  1042
Ser Arg Pro Phe Val Cys Glu His Ala Gly Cys Gly Lys Thr Phe Ala
    275                 280                 285

atg aaa caa agt ctc act agg cat gct gtt gta cat gat cct gac aag  1090
Met Lys Gln Ser Leu Thr Arg His Ala Val Val His Asp Pro Asp Lys
290                 295                 300                 305

aag aaa atg aag ctc aaa gtc aaa aaa tct cgt gaa aaa cgg agt ttg  1138
Lys Lys Met Lys Leu Lys Val Lys Lys Ser Arg Glu Lys Arg Ser Leu
```

```
                310                 315                 320

gcc tct cat ctc agt gga tat atc cct ccc aaa agg aaa caa ggg caa   1186
Ala Ser His Leu Ser Gly Tyr Ile Pro Pro Lys Arg Lys Gln Gly Gln
            325                 330                 335

ggc tta tct ttg tgt caa aac gga gag tca ccc aac tgt gtg gaa gac   1234
Gly Leu Ser Leu Cys Gln Asn Gly Glu Ser Pro Asn Cys Val Glu Asp
        340                 345                 350

aag atg ctc tcg aca gtt gca gta ctt acc ctt ggc taa               1273
Lys Met Leu Ser Thr Val Ala Val Leu Thr Leu Gly
    355                 360                 365


<210> 2
<211> 365
<212> PRT
<213> Human

<400> 2
Met Asp Pro Pro Ala Val Val Ala Glu Ser Val Ser Ser Leu Thr Ile
  1               5                  10                  15

Ala Asp Ala Phe Ile Ala Ala Gly Glu Ser Ser Ala Pro Thr Pro Pro
             20                  25                  30

Arg Pro Ala Leu Pro Arg Arg Phe Ile Cys Ser Phe Pro Asp Cys Ser
         35                  40                  45

Ala Asn Tyr Ser Lys Ala Trp Lys Leu Asp Ala His Leu Cys Lys His
     50                  55                  60

Thr Gly Glu Arg Pro Phe Val Cys Asp Tyr Glu Gly Cys Gly Lys Ala
 65                  70                  75                  80

Phe Ile Arg Asp Tyr His Leu Ser Arg His Ile Leu Thr His Thr Gly
                 85                  90                  95

Glu Lys Pro Phe Val Cys Ala Ala Thr Gly Cys Asp Gln Lys Phe Asn
            100                 105                 110

Thr Lys Ser Asn Leu Lys Lys His Phe Glu Arg Lys His Glu Asn Gln
        115                 120                 125

Gln Lys Gln Tyr Ile Cys Ser Phe Glu Asp Cys Lys Lys Thr Phe Lys
    130                 135                 140

Lys His Gln Gln Leu Lys Ile His Gln Cys Gln His Thr Asn Glu Pro
145                 150                 155                 160

Leu Phe Lys Cys Thr Gln Glu Gly Cys Gly Lys His Phe Ala Ser Pro
                165                 170                 175

Ser Lys Leu Lys Arg His Ala Lys Ala His Glu Gly Tyr Val Cys Gln
            180                 185                 190
```

Lys Gly Cys Ser Phe Val Ala Lys Thr Trp Thr Glu Leu Leu Lys His
195 200 205

Val Arg Glu Thr His Lys Glu Glu Ile Leu Cys Glu Val Cys Arg Lys
210 215 220

Thr Phe Lys Arg Lys Asp Tyr Leu Lys Gln His Met Lys Thr His Ala
225 230 235 240

Pro Glu Arg Asp Val Cys Arg Cys Pro Arg Glu Gly Cys Gly Arg Thr
245 250 255

Tyr Thr Thr Val Phe Asn Leu Gln Ser His Ile Leu Ser Phe His Glu
260 265 270

Glu Ser Arg Pro Phe Val Cys Glu His Ala Gly Cys Gly Lys Thr Phe
275 280 285

Ala Met Lys Gln Ser Leu Thr Arg His Ala Val Val His Asp Pro Asp
290 295 300

Lys Lys Lys Met Lys Leu Lys Val Lys Lys Ser Arg Glu Lys Arg Ser
305 310 315 320

Leu Ala Ser His Leu Ser Gly Tyr Ile Pro Pro Lys Arg Lys Gln Gly
325 330 335

Gln Gly Leu Ser Leu Cys Gln Asn Gly Glu Ser Pro Asn Cys Val Glu
340 345 350

Asp Lys Met Leu Ser Thr Val Ala Val Leu Thr Leu Gly
355 360 365

<210> 3
<211> 1273
<212> DNA
<213> Human

<400> 3
atgcgcagca gcggcgccga cgcggggcgg tgcctggtga ccgcgcgcgc tcccggaagt 60

gtgccggcgt cgcgcgaagg ttcagcaggg agccgtgggc cgggcgcgcc ggttcccggc 120

acgtgtctcg gcacgtggca gcgcgcctgg ccctgggctt ggaggcgccg gcgccctgga 180

tccgccggcc gtggtcgccg agtcggtgtc gtccttgacc atcgccgacg cgttcattgc 240

agccggcgag agctcagctc cgaccccgcc gcgccccgcg cttcccagga ggttcatctg 300

ctccttccct gactgcagcg ccaattacag caaagcctgg aagcttgacg cgcacctgtg 360

caagcacacg ggggagagac catttgtttg tgactatgaa gggtgtggca aggccttcat 420

cagggactac catctgagcc gccacattct gactcacaca ggagaaaagc cgtttgtttg 480

tgcagccact ggctgtgatc aaaaattcaa cacaaaatca aacttgaaga aacattttga 540

acgcaaacat gaaaatcaac aaaaacaata tatatgcagt tttgaagact gtaagaagac 600

ctttaagaaa catcagcagc tgaaaatcca tcagtgccag cataccaatg aacctctatt 660

caagtgtacc caggaaggat gtgggaaaca ctttgcatca cccagcaagc tgaaacgaca 720

tgccaaggcc cacgagggct atgtatgtca aaaaggatgt tcctttgtgg caaaaacatg 780

gacggaactt ctgaaacatg tgagagaaac ccataaagag gaaatactat gtgaagtatg 840

ccggaaaaca tttaaacgca aagattacct taagcaacac atgaaaactc atgccccaga 900

aagggatgta tgtcgctgtc caagagaagg ctgtggaaga acctatacta ctgtgtttaa 960

tctccaaagc catatcctct ccttccatga ggaaagccgc ccttttgtgt gtgaacatgc 1020

tggctgtggc aaaacatttg caatgaaaca aagtctcact aggcatgctg ttgtacatga 1080

tcctgacaag aagaaaatga agctcaaagt caaaaaatct cgtgaaaaac ggagtttggc 1140

ctctcatctc agtggatata tccctcccaa aaggaaacaa gggcaaggct tatctttgtg 1200

tcaaaacgga gagtcaccca actgtgtgga agacaagatg ctctcgacag ttgcagtact 1260

tacccttggc taa 1273

<210> 4
<211> 1213
<212> DNA
<213> Human

<400> 4
gtgccggcgc cgcgcgaagg ttcagcaggg agccgtgggc cgggcgcgcc ggttcccggc 60

acgtgtctcg gcacgtggca gcgcgcctgg ccctgggctt ggaggcgccg gcgccctgga 120

tccgccggcc gtggtcgccg agtcggtgtc gtccttgacc atcgccgacg cgttcattgc 180

agccggcgag agctcagctc cgaccccgcc gcgccccgcg cttcccagga ggttcatctg 240

ctccttccct gactgcagcg ccaattacag caaagcctgg aagcttgacg cgcacctgtg 300

caagcacacg ggggagagac catttgtttg tgactatgaa gggtgtggca aggccttcat 360

cagggactac catctgagcc gccacattct gactcacaca ggagaaaagc cgtttgtttg 420

tgcagccact ggctgtgatc aaaaattcaa cacaaaatca aacttgaaga aacattttga 480

acgcaaacat gaaaatcaac aaaaacaata tatatgcagt tttgaagact gtaagaagac 540

ctttaagaaa catcagcagc tgaaaatcca tcagtgccag cataccaatg aacctctatt 600

caagtgtacc caggaaggat gtgggaaaca ctttgcatca cccagcaagc tgaaacgaca 660

tgccaaggcc cacgagggct atgtatgtca aaaaggatgt tcctttgtgg caaaaacatg 720

gacggaactt ctgaaacatg tgagagaaac ccataaagag gaaatactat gtgaagtatg 780

ccggaaaaca tttaaacgca aagattacct taagcaacac atgaaaactc atgccccaga 840

aagggatgta tgtcgctgtc caagagaagg ctgtggaaga acctatacta ctgtgtttaa 900

tctccaaagc catatcctct ccttccatga ggaaagccgc ccttttgtgt gtgaacatgc 960

tggctgtggc aaaacatttg caatgaaaca aagtctcact aggcatgctg ttgtacatga 1020

tcctgacaag aagaaaatga agctcaaagt caaaaaatct cgtgaaaaac ggagtttggc 1080

ctctcatctc agtggatata tccctcccaa aaggaaacaa gggcaaggct tatctttgtg 1140

tcaaaacgga gagtcaccca actgtgtgga agacaagatg ctctcgacag ttgcagtact 1200

taccettggc taa 1213

<210> 5
<211> 34
<212> DNA
<213> Human

<400> 5
cggggtacca aaaatgcgca gcagcggcgc cgac 34

<210> 6
<211> 21
<212> DNA
<213> Human

<400> 6
tccttccctg actgcagcgc c 21

<210> 7
<211> 20
<212> DNA
<213> Human

<400> 7
tgcacaggtg cgcgtcaagc 20

<210> 8
<211> 20
<212> DNA
<213> Human

<400> 8
cacaaacaaa tggtctctcc 20

<210> 9

<211> 30
<212> DNA
<213> Human

<400> 9
cggtctagat tagccaaggg taagtactgc 30

<210> 10
<211> 30
<212> DNA
<213> Human

B7

<400> 10
cctcccgggg ccaagggtaa gtactgcaac 30